DSCOVR EPIC Vegetation Earth System Data Record

Science Data Product Guide

Maturity level: Provisional

Y. Knyazikhin, W. Song, B. Yang, T. Park and R.B. Myneni

Boston University

Boston, April 22, 2018

Distributed by the Atmospheric Science Data Center
http://eosweb.larc.nasa.gov

TABLE OF CONTENTS

1. INTRODUCTION

1.1. Purpose. This document describes Level 2 Vegetation Earth System Data Record (VESDR) derived from the Earth Polychromatic Imaging Camera (EPIC) onboard the Deep Space Climate Observatory (DSCOVR). It provides file structure for the geophysical and ancillary science data products. The VESDR parameters are summarized in Table 1.

Table 1: Vegetation Parameter Suite in the Level 2 Vegetation Earth System Data Record (VESDR) Product

Parameter name	Units	Resolution		Comments
		Temporal	Spatial	
Normalized Difference Vegetation Index (NDVI)	none	65 to 110 min	10 km	difference between Bidirectional Reflectance Factor (BRF) at 779.5 nm and 680 nm normalized by their sum
Fraction vegetation absorbed Photosynthetically Active Radiation (FPAR)	fraction	65 to 110 min	10 km	fraction of photosynthetically active radiation (400 – 700nm) absorbed by vegetation
Leaf Area Index (LAI)	$\frac{\mathrm{m}^2_{\mathrm{plant}}}{\mathrm{m}^2_{\mathrm{ground}}}$	65 to 110 min	10 km	one-sided green leaf area per unit ground area in broadleaf canopies and the projected needle area in coniferous canopies
Sunlit Leaf Area Index(SLAI)	$\frac{m^2_{\mathrm{sunlit}}}{m^2_{\mathrm{ground}}}$	65 to 110 min	10 km	one-sided sunlit green leaf area per unit ground area in broadleaf canopies and the projected sunlit needle area in coniferous canopies
Precision of Leaf Area Index (Dlai)	$\frac{m^2_{\mathrm{plant}}}{m^2_{\mathrm{gound}}}$	65 to 110 min	10 km	retrieval dispersion of LAI
Directional Area Scattering Factor (DASF)	none	65 to 110 min	10 km	Estimate of Canopy Bidirectional Reflectance Factor as if the foliage does not absorb radiation
Quality Assessment variable(QA_VESDR)	none	65 to 110 min	10 km	Overall quality of the VESDR parameters

With the exception of LAI, all VESDR parameters vary with the sun-sensor geometry. The VESDR file also includes Solar Zenith Angle (SZA), Solar Azimuthal Angle (SAA), View Zenith (VZA) and Azimuthal (VAA) angles at the same temporal and spatial resolutions (Sect. 3).

The DSCOVR EPIC Science Algorithm Team also provides two ancillary science data products, namely, *10 km Land Cover Type* and *Distribution of Land Cover Types within 10 km EPIC pixel*. The products were derived from 500m MODIS land cover type 3 product (MCDLCHKM), which was generated from 2008, 2009 and 2010 land cover products (MCD12Q1, v051). The ancillary data sets are summarized in Table 2.

All products are projected on 10 km sinusoidal (SIN) grid and written in the standard Hierarchical Data Format 5 (HDF5) using HDF-defined data models (http://www.hdfgroup.org/HDF5/). The EPIC VESDR and ancillary data products are publicly available from the NASA Langley Atmospheric Science Data Center (https://eosweb.larc.nasa.gov/project/dscovr/dscovr_table).

Table 2: Ancillary science data product derived form 500m MODIS land cover type 3 product

Parameter name	Units	Resolution		Comments
		Temporal	Spatial	
Land Cover Type	none	static	10 km	10 km SIN Land Cover type
Land Cover Type Distribution	none	static	10 km	Distribution of land cover types within 10 km EPIC pixel

1.2. Product maturity level. Definitions of product maturity levels developed by the MISR team are adopted (https://www-misr.jpl.nasa.gov/getData/maturityLevels/). The DSCOVR EPIC VESDR product is released at Provisional quality level, i.e.,

- Incremental improvements are still occurring. Obvious artifacts or blunders observed in prerelease product have been identified and either minimized or documented
- General research community is encouraged to participate in the quality assessment and validation, but need to be aware that product validation and quality assessment are ongoing
- Parameter may be used in publications as long as provisional quality is indicated by the authors. Users are urged to contact science team representatives prior to use of the data in publications, and to recommend members of the instrument teams as reviewers
- The Data Quality Summary states estimated uncertainties
- May be replaced in the archive when an upgraded product becomes available, but should be reproducible upon demand

DSCOVR EPIC data products begin in a provisional state, and advance through a series of maturity levels, from Provisional to Validated status, i.e. from a developmental status to a scientifically proven status.

1.3. DSCOVR EPIC documents. Project documents are available at https://eosweb.larc.nasa.gov/project/dscovr/dscovr_table. DSCOVR EPIC publications can be found at https://epic.gsfc.nasa.gov/science/pubs. The VESDR theoretical basis was documents in

[1] Yang, B., Knyazikhin, Y., Mõttus, M., Rautiainen, M., Stenberg, P., Yan, L., Chen, C., Yan, K., Choi, S., Park, T., & Myneni, R.B. (2017). Estimation of leaf area index and its sunlit portion from DSCOVR EPIC data: Theoretical basis. *Remote Sensing of Environment*, 198, 69-84.doi: /10.1016/j.rse.2017.05.033

An overview of the DSCOVR EPIC project is documented in

[2] Marshak, A., Herman, J., Szabo, A., Blank, K., Cede, A., Carn, S., Geogdzhayev, I., Huang, D., Huang, L.-K., Knyazikhin, Y., Kowalewski, M., Krotkov, N., Lyapustin, A., McPeters, R., Torres, O., & Yang, Y. Earth Observations from DSCOVR/EPIC Instrument. *Bulletin of the American Meteorological Society*, doi:/10.1175/BAMS-D-17-0223.1.

The Directional Area Scattering Factor (DASF) is a new structural parameter that estimates the canopy BRF if the leaves do not absorb radiation. Its definition and analysis of its value for remote sensing of leaf biochemistry can be found in

[3] Knyazikhin, Y., Schull, M.A., Stenberg, P., Mõttus, M., Rautiainen, M., Yang, Y., Marshak, A., Latorre Carmona, P., Kaufmann, R.K., Lewis, P., Disney, M.I., Vanderbilt, V., Davis, A.B., Baret, F., Jacquemoud, S., Lyapustin, A., & Myneni, R.B. (2013). Hyperspectral remote sensing of foliar nitrogen content. *Proceedings of the National Academy of Sciences*, 110, E185-E192

1.4. Revisions. This is the first version of the document. It can be downloaded, distributed, and cited. Revisions of the Science Data Product Guide will be detailed in this section.

2. EXPERIMENT OVERVIEW

The Deep Space Climate Observatory (DSCOVR) mission is a multiagency (National Oceanic and Atmospheric Administration [NOAA], U.S. Air Force, and NASA) mission launched from Cape Canaveral, Florida on February 11, 2015 with the primary goal of making unique space weather measurements from the first Sun-Earth Lagrange point (L1). The L1 point is on the direct line between Earth and the Sun located 1.5 million km sunward from Earth. The spacecraft is orbiting this point in a six month Lissajous orbit with a Sun-Earth-View (SEV) angle varying between 4.5^o and 11.5^o. The primary science objective of the DSCOVR mission is to provide solar wind thermal plasma and magnetic field measurements to enable space weather forecasting by NOAA.

The DSCOVR hosts NASA Earth-Observing Instrument, the Earth Polychromatic Imaging Camera (EPIC). The EPIC provides measurements of the radiation reflected by Earth in ten wavelengths and images of the sunlit side of Earth for science applications.

2.1. EPIC instrument characteristics. The EPIC instrument collects multispectral data of the Earth in ten wavelengths. The spectral band characteristics are summarized in Table 3.

Table 3: EPIC spectral band composition

Wavelength, nm	FWHM,nm	Nominal Product
317.5±0.1	1±0.2	Ozon
325±0.1	2±0.2	Ozon
340±0.3	3±0.6	Ozon, Aerosols, Clouds
388±0.3	3±0.6	Aerosols, Clouds
443±1	3±0.6	Aerosols
551±1	3±0.6	Aerosols, Vegetation
680±0.2	2±0.4	Aerosols, Vegetation, Clouds,O_2 B-Band Reference
687.75±0.2	0.8±0.2	O_2 B-Band Cloud Height
764±0.2	1±0.2	O_2 A-Band Cloud Height, Aerosol Height
779±0.3	2±0.4	O_2 A-Band Reference, Vegetation

2.2. Rationale for the DSCOVR EPIC VESDR product. Fraction vegetation absorbed Photosynthetically Active Radiation (FPAR), Leaf Area Index (LAI), its sunlit counterpart (SLAI), and Normalized Difference Vegetation Index (NDVI are useful for (a) monitoring variability and change in global vegetation due to climate and anthropogenic influences, (b) modeling climate, carbon and water cycles, and (c) improving forecasting of near surface weather. The Directional Area Scattering Factor provides information critical to accounting for structural contributions to measurements of leaf biochemistry from remote sensing. Whereas LAI is a standard product of many satellite missions, global diurnal courses of FPAR, NDVI, SLAI and DASF are new satellite derived products.

3. VESDR SUN-SENSOR GEOMETRY

The sun-sensor geometry is expressed in a right-handed coordinate system in which the Z-axis (shown as "+Z" in Fig. 1) is aligned with the normal to the surface reference ellipsoid (defined by the World Geodetic System 1984, WGS84), and points toward the center of the Earth. The X-axis is aligned with a great circle and points toward the north pole. The Y-axis is orthogonal to both of them.

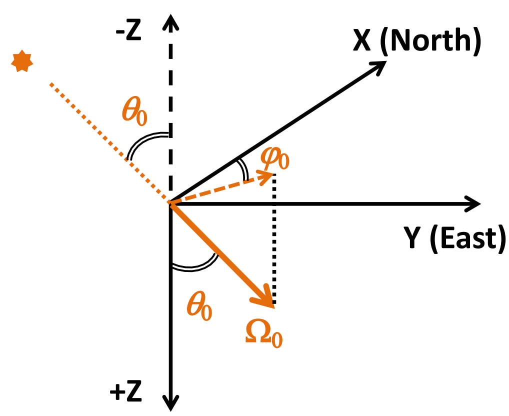

Figure 1. Right-handed coordinate system in which the Z-axis (shown as "+Z") points toward the center of the Earth. The X-axis and Y-axis point toward the North and East, respectively. The direction (unit vector) Ω_0 has an azimuthal angle, φ_0, measured clockwise from the local north vector (X) to the projection of Ω_0 onto the XY plane, and a polar angle, θ_0, with respect to the +Z axis.

The Sun and sensor positions are represented by unit vectors Ω_0 and Ω directed *downward* from the *Sun to target* (i.e., point at the Earth surface) and from *sensor to target*, respectively. Their polar (θ_0 and θ) and azimuthal (φ_0 and φ) angles are given in the right-handed coordinate system (Fig. 1). Their ranges are between 0 and 90° (polar angles) and between 0 and 360° (azimuthal angles). This coordinate system is inherited from the upstream DSCOVER EPIC L2 MAIAC surface reflectance product, which is input to the VESDR retrieval algorithm. In this coordinate system the Solar Zenith Angle (SZA, the angle between the target-to-Sun direction, $-\Omega_0$, and the –Z axis) coincides with the polar angle of Ω_0, i.e., $\theta_0 = SZA$.

The Earth-observing geometry of the EPIC instrument is characterized by a nearly constant phase angle (the angle between directions to the Sun and to the sensor) between 4.5° and 11.5°. The phase angle, γ, can be calculated as

$$\gamma = \operatorname{acos}(\cos\theta\cos\theta_0 + \sin\theta\sin\theta_0\cos(\varphi - \varphi_0)). \quad (1)$$

4. PRODUCT TILING

The VESDR and ancillary science data products are projected on 10 km sinusoidal grid. The globe is divided into 4 horizontal tiles along the east-west, and 2 vertical tiles along the north-south axes (Fig. 2). Each tile is identified by its horizontal (from 0 to 3) and vertical (from 0 to 1) coordinates, e.g., tile01. Dimension of one tile is 1000x1002.

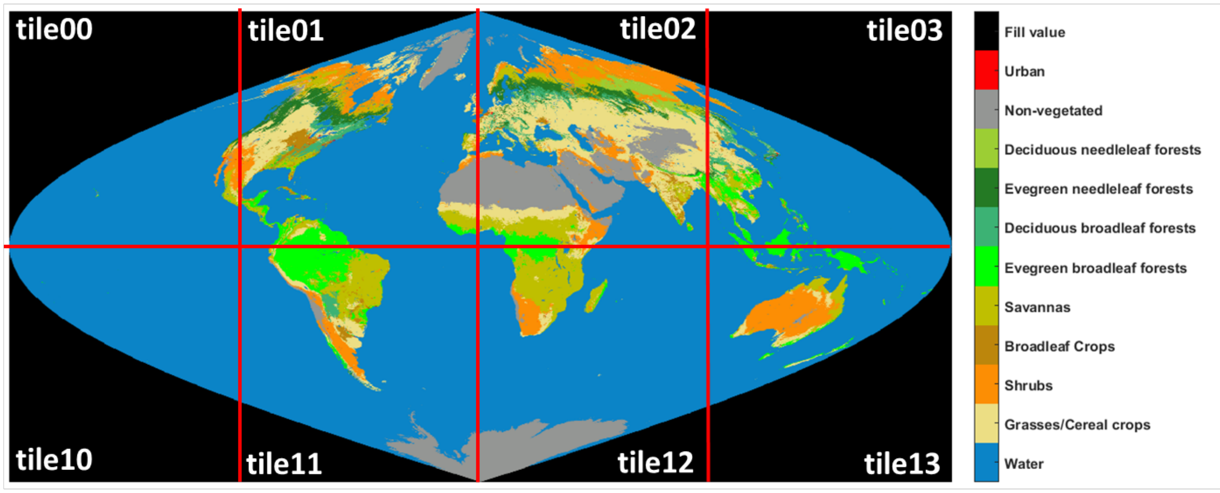

Figure 2. 10 km SIN DSCOVR EPIC land cover type. The globe is divided into 8 equal tiles. Each tile is identified by its horizontal (from 0 to 3) and vertical (from 0 to 1) coordinates. The VESDR and ancillary science data products use this tiling structure.

5. LEVEL 2 VESDR PRODUCT

5.1. VESDR product file name

The file name containing VESDR parameters is

DSCOVR_EPIC_L2_VESDR_V1_YYYYMMDDHHMMSS_V2.h5

Here V1 and V2 are versions of the VESDR product and L2B TOA reflectance data, respectively. Current versions are V1=01 and V2=02. YYYYMMDDHHMMSS signifies date and GMT time of EPIC image acquisition. For example, file DSCOVR_EPIC_L2_VESDR_01_20160823141930_02.h5 contains VESDR parameters for an EPIC image acquired on August 23, 2016 (20160823) at 14h10m30s GMT (141930).

5.2. HDF file structure

The VESDR product is distributed as standard Hierarchical Data Format 5 (HDF5) file. The data are compressed using the lossless gzip option provided by the HDF5 FORTRAN Application Programming Interface (API). Compression level is 4. On average L2 VESDR product is about 20-22 megabytes (MB).

In the HDF5 file, data are grouped by tiles. Each group contains geophysical parameters, associated quality assessment variables (QA_VESDR and Dlai) and sun-sensor geometry. The root group directory contains a set of attributes that describes the content of the HDF5 file. Figure 3 illustrates a snapshot of the layout of a L2 VESDR product.

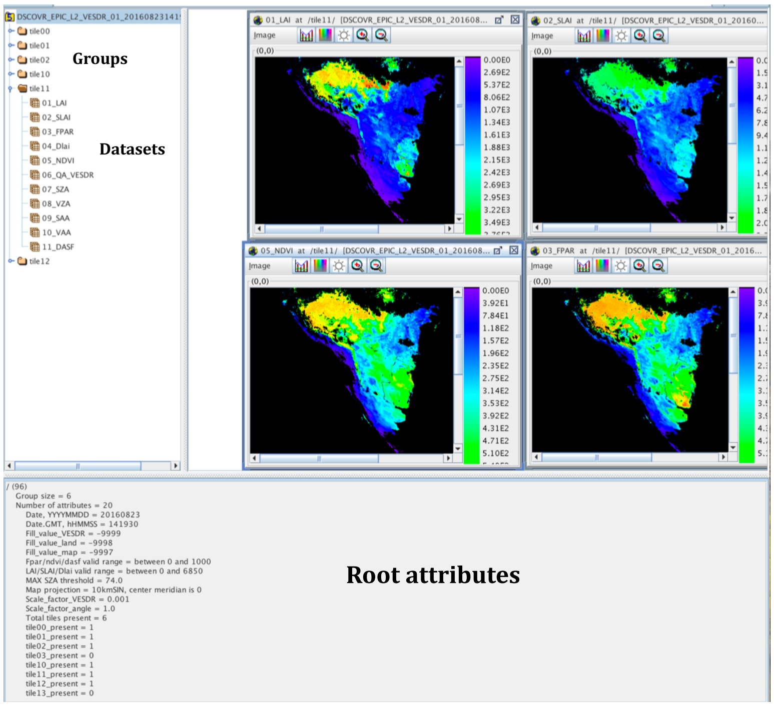

Figure 3. Structure of the L2 VESDR product. Data are grouped by tiles. Each group contains geophysical paramet-ers, associated quality assessment variables and sun-sensor geometry. The root attribute directory provides general information about the VESDR product.

5.3. Root attributes

Root attributes include date and time of the EPIC image acquisition, fill values, parameter's valid ranges, map projection, scale factors and list of tiles present in the file. Details are summarized in Table 4.

Table 4: L2 VESDR root attributes

Attribute name	Value, range	Type	Description
Date, YYYYMMDD	20160613-current date	32 bit integer	date of EPIC image acquisition
Date.GMT, hHMMSS	00000 - 235959	32 bit integer	GMT of EPIC image acquisition
Fill_value_VESDR	-9999	16 bit integer	VESDR parameter was not generated
Fill_value_land	-9998	16 bit integer	non-vegetated pixel
Fill_value_map	-9997	16 bit integer	out of map boundary ("black area" in Fig. 2)
Fpar/ndvi/dasf valid range	0-1000	string	valid range of FPAR, NDVI and DASF
LAI/SLAI/Dlai valid range	0-6850	string	valid range of LAI, SLAI and dLAI
Max SZA threshold	74.0	32 bit floating point	VESDR algorithm does not process pixel if the SZA exceeds Max_SZA_threshold
Map projection	10 km SIN, center meridian is 0	string	VESDR parameters are projected on 10 km sinusoidal grid

Scale_factor_VESDR	0.001	32 bit floating point	VESDR parameter should be multiplied by the scale factor to convert its DN value to physical value
Scale_factor_angle	1.0	32 bit floating point	Sun-sensor geometry parameters should be multiplied by the scale factor to get their physical value
Total tiles present	1-8	8 bit integer	Total number of tiles present in the VESDR file
tile00_present	0,1	8 bit integer	Indicates if tile00 present (value=1) in dataset
tile01_present	0,1	8 bit integer	Indicates if tile01 present (value=1) in dataset
tile02_present	0,1	8 bit integer	Indicates if tile02 present (value=1) in dataset
tile03_present	0,1	8 bit integer	Indicates if tile03 present (value=1) in dataset
tile10_present	0,1	8 bit integer	Indicates if tile10 present (value=1) in data et
tile11_present	0,1	8 bit integer	Indicates if tile11 present (value=1) in dataset
tile12_present	0,1	8 bit integer	Indicates if tile12 present (value=1) in data et
tile13_present	0,1	8 bit integer	Indicates if tile13 present (value=1) in dataset

5.4. Datasets

Each group contains geophysical parameters, associated quality assessment variables (QA_VESDR and Dlai) and sun-sensor geometry. Description of the datasets is given in Table 5.

Table 5: L2 VESDR datasets

Name of dataset	Valid range	Data type	Description
01_LAI	0-6850	16 bit integer	Leaf Area Index
02_SLAI	0-6850	16 bit integer	Sunlit Leaf Area Index
03_FPAR	0-1000	16 bit integer	fraction of photosynthetically active radiation (400 – 700nm) absorbed by vegetation
04_Dlai	0-6850	16 bit integer	Precision of Leaf Area Index
05_NDVI	0-1000	16 bit integer	Normalized Difference Vegetation Index
06_QA_VESDR	0-767	16 bit integer	Quality Assessment variable. See section 5.5
07_SZA	0-90	32 bit floating point	Polar angle (in DEG) of the Sun-to-target direction as defined in Sect. 3
08_VZA	0-90	32 bit floating point	Polar angle (in DEG) of the sensor-to-target direction as defined in Sect. 3
09_SAA	0-360	32 bit floating point	Azimuthal angle (in DEG) of the Sun-to-target direction as defined in Sect. 3
10_VAA	0-360	32 bit floating point	Azimuthal angle (in DEG) of the sensor-to-target direction as defined in Sect. 3
11_DASF	0-1000	16 bit integer	Estimate of Canopy Bidirectional Reflectance Factor as if the foliage does not absorb radiation

5.5. Quality assessment dataset

5.5.1. Information content of QA dataset. Quality assessment variable, 06_QA_VESDR, includes quality control information on VESDR algorithm performance (bits 0 to 5) and Status_QA (bits 6 to 9). The latter is provided by the upstream DSCOVR EPIC L2 MAIAC surface reflectance product. The DSCOVR EPIC MAIAC product is input to the VESDR retrieval technique. 06_QA_VESDR therefore provides information about quality of both input to the VESDR algorithm and the VESDR algorithm output. Figure 4 shows structure of 06_QA_VESDR. Details are given in Table 6.

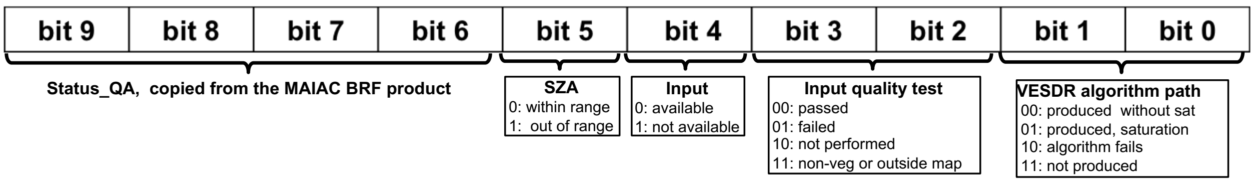

Figure 4. Information content of 06_QA_VESDR

Table 5: Values of QA_VESDR

QA name	Bits	Binary value	Decimal value	Description
VESDR algorithm path	0-1	00	0	VESDR parameters produced. No saturation
		01	1	VESDR parameters produced under a saturation condition
		10	2	VESDR algorithm failed to generate parameters
		11	3	VESDR parameters were not produced. Other reasons. See bits 2-5
Input quality test	2-3	00	0	Input quality test passed
		01	1	Input quality test failed
		10	2	Input quality test was not performed because BRF at NIR and/or Green spectral bands were not available
		11	3	VESDR parameters were not produced because pixel was not-vegetated or out of map. Bits 0-5 are set to 1 in this case
Input availability	4	0	0	BRF at NIR and Red spectral bands were available
		1	1	Parameters were not produced because BRFs at NIR and/or Red spectral bands were not available.
SZA	5	0	0	SZA is between 0º and maxSZAthreshod. See root attributes.
		1	1	SZA is outside of the acceptable range. Parameters not produced
Status_QA	6-9	0000	0	No clouds, CM_CLEAR_WATER
		0001	1	No clouds, CM_CLEAR_WATERSED
		0010	2	1 neighbor cloud
		0011	3	>1 neighbor clouds
		0100	4	no retrieval (cloudy, or whatever)
		0101	5	definition is not provided
		0110	6	for H>3.5km, no retrieval
		0111	7	definition is not provided
		0100	8	sun glint
		1001	9	land-water misclassified
		1010	10	CoxMunk too high
		1011	11	info not available

5.5.2. Saturation, Retrieval Index and input quality test. In the case of dense canopies, the reflectances saturate and therefore are weakly sensitive to changes in canopy properties. The reliability of parameters retrieved under the condition of saturation is low. Such retrievals are flagged by setting decimal value of the VESDR_algorithm_path to 1 (binary value ='01').

The retrieval index, RI, is the percentage of pixels with valid BRF for which the VESDR algorithm produced a retrieval. The index characterizes the spatial coverage of the geophysical parameters. This important characteristic of the algorithm performance can be calculated as

$$RI = \frac{N(\text{bits01} =' 00' \text{or bits01} =' 01')}{N(\text{bit4} =' 0')}. \quad (2)$$

Here the numerator represents number of pixels for which the VESDR_algorithm_path is 0 or 1. The denominator is the number of pixels for which a value of bit 4 is 0.

For vegetated pixels at weakly absorbing wavelengths, the BRF to leaf albedo ratio is linearly related to BRF, i.e.,

$$\frac{BRF_\lambda}{\omega_\lambda} = pBRF_\lambda + R\,, \quad (3)$$

where the slope, p, and intercept, R, are the recollision probability and escape factor. We use BRF at green and NIR EPIC bands to estimate the slope, p, of a line passing points $\left(\frac{BRF_{green}}{\omega_{green}}, BRF_{green}\right)$ and $\left(\frac{BRF_{NIR}}{\omega_{NIR}}, BRF_{NIR}\right)$ on the $\frac{BRF}{\omega}$ vs BRF plane. Its value is given by

$$p = \frac{\frac{BRF_{NIR}}{\omega_{NIR}} - \frac{BRF_{green}}{\omega_{green}}}{BRF_{NIR} - BRF_{green}}. \quad (4)$$

Here ω_λ represents a fixed leaf albedo at NIR and green spectral bands. Its values at these bands are set to 0.4898 (green) and 0.9789 (NIR) in the VESDR operational algorithm. Our analyses suggest that Eq. (4) takes values between 0 and 1 only for vegetated surfaces. For BRF at green and NIR spectral bands over non-vegetated land, water or cloud-contaminated pixels, Eq.(4) generates values outside of the 0 to 1 range. This property underlies the input quality test: bits 2-3 are set to '00' if p is between 0 and 1, and to '01', otherwise. The VESDR algorithm processes pixels irrespective of the test result. The input_quality_test QA is just a warning that VESDR parameters were retrieved using input BRF of suspicious quality.

Figure 4 shows distribution of $\alpha = \text{atan}(p)$, $0° \le \alpha \le 180°$ derived from EPIC L1B TOA reflectance data. One can see that values of α corresponding to cloud free land, vegetation, ocean and cloud contaminated pixels tend to occupy different spaces within the 0^o to 180^o interval.

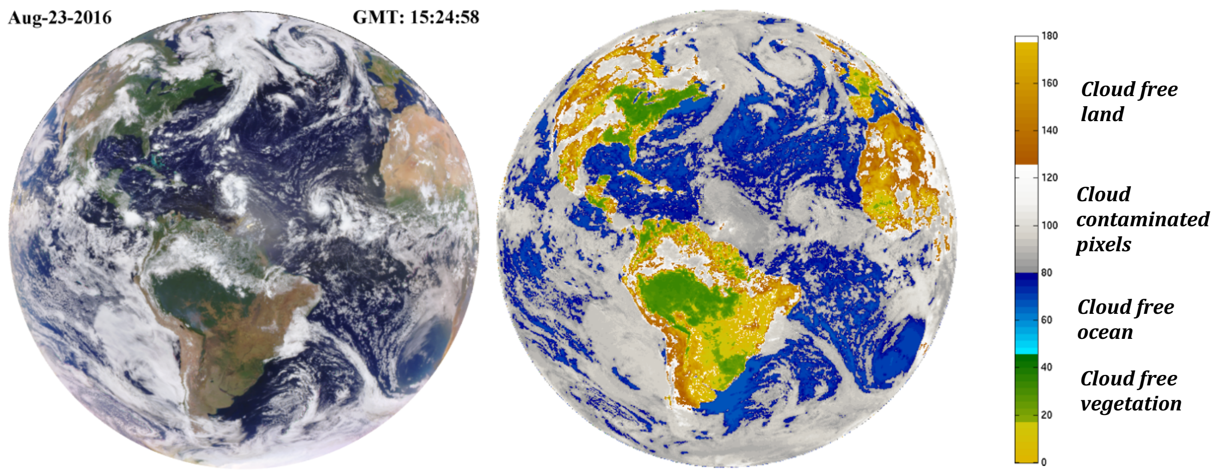

Figure 4. Left panel shows an EPIC RGB image taken on August 23, 2016 at 15:24:58GMT. Eq. (4) was applied to each image pixel. The slope p was converted to angle, $\alpha = \text{atan}(p)$, $\alpha \in [0°, 180°]$, between the line and BRF axis. Right panel shows distribution of α over the EPIC image. Its values corresponding to cloud free land, vegetation, ocean and cloud contaminated pixels tend to occupy different spaces within the 0^o to 180^o interval.

6. ANCILLARY SCIENCE DATA PRODUCTS

A land cover map is an important ancillary data layer used by the VESDR retrieval algorithm. The global classification of canopy structural types utilized in the Collection 6 MODIS LAI/FPAR algorithm is adopted. Global vegetation is stratified into eight canopy architectural types, or biomes. The eight biomes are Grasses and Cereal Crops (B1), Shrubs (B2), Broadleaf Crops (B3), Savannas (B4), Evergreen Broadleaf Forests (B5), Deciduous Broadleaf Forests (B6), Evergreen Needle Leaf Forests (B7) and Deciduous Needle Leaf Forests (B8).

The VESDR ancillary science data products include *10 km Land Cover Type* and *Distribution of Land Cover Types within 10 km EPIC pixel*. These products were derived from the MODIS 8-biome SIN 500 m resolution land cover type 3 product (MCDLCHKM), which was generated from 2008, 2009 and 2010 MODIS land cover products (MCD12Q1, v051).

6.1. DSCOVR EPIC land cover type

The MODIS Land Cover Product is projected on 500 m sinusoidal (SIN) grid. A 10 km EPIC SIN grid pixel therefore contains about 400 MODIS pixels with known land cover types. The EPIC land cover type is assigned based on the dominant land cover fraction. If there are several land cover types with equal frequency, biome type with highest biome number BN is taken as the EPIC land cover type. For example, if B5 (Deciduous Broadleaf Forests), B4 (Shrubs) and B1 (Grasses and Cereal Crops) occupy 40%, 40% and 20% of the pixel area, then B5 is assigned to the EPIC land cover type. The most frequent land cover type numbers are also stored in the DSCOVR EPIC land cover file. Figure 2 shows 10 km SIN DSCOVR EPIC land cover type.

DSCOVR EPIC land cover product file name is DSCOVR_EPIC_ANC_LCTYPE_MCD12Q1_51.h5.

6.1.1. HDF file structure. In the HDF5 file, data are grouped by tiles. Each root group contains two data sets, "Land_Cover_Type_3" and "Multi_Land_Cover_Types_presented," as well as a tile group "Geolocation" with two datasets, "Latitude" and "Longitude." The root group directory contains a set of attributes that describes the content of the HDF5 file.

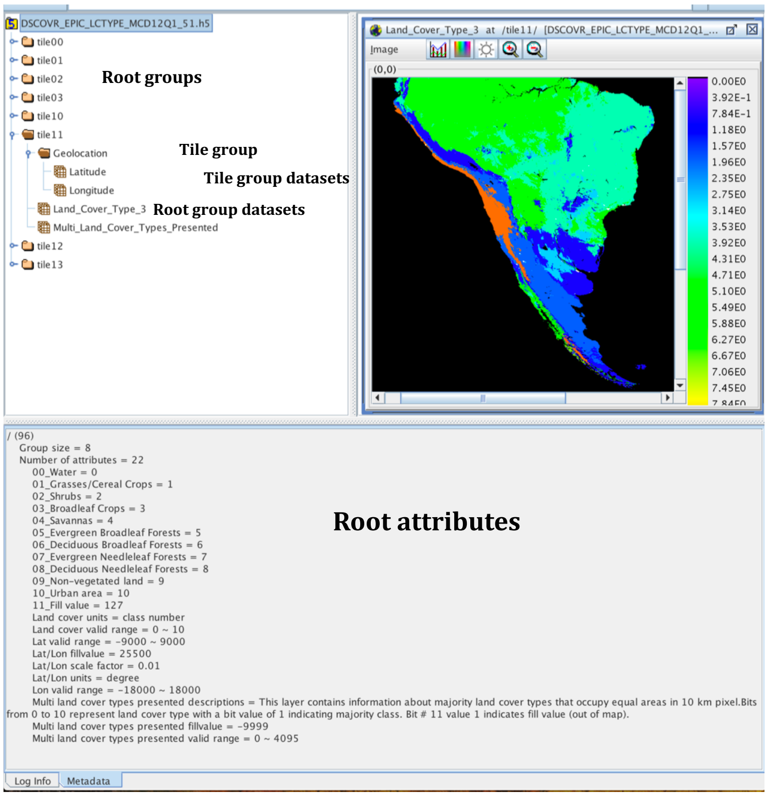

Figure 5. Structure of the ancillary 10 km EPIC Land Cover Type Product

6.1.2. Root attributes. Root attributes include land cover type IDs and associated values, fill values, parameter valid ranges and short description of the "Multi_Land_Cover_Types_presented" dataset. Details are summarized in Table 6.

Table 6: Land Cover type root attributes

Attribute name	Value	Type	Description
00_Water	0	8 bit integer	Pixel is classified as water
01_Grasses/Cereal Crops	1	8 bit integer	Pixel is classified as B1: Grasses and Cereal Crops
02_Shrubs	2	8 bit integer	Pixel is classified as B2: Shrubs
03_Broadleaf Crops	3	8 bit integer	Pixel is classified as B3: Broadleaf Crops
04_Savannas	4	8 bit integer	Pixel is classified as B4: Savannas
05_Evergreen Broadleaf Forests	5	8 bit integer	Pixel is classified as B5: Evergreen Broadleaf Forests
06_Deciduous Broadleaf Forests	6	8 bit integer	Pixel is classified as B6 : Deciduous Broadleaf Forests
07_Evergreen Needle Leaf Forests	7	8 bit integer	Pixel is classified as B7: Evergreen Needle Leaf Forests
08_ Deciduous Needle Leaf Forests	8	8 bit integer	Pixel is classified as B8: Deciduous Needle Leaf Forests
09_Non-vegetated land	9	8 bit integer	Pixel is classified as non-vegetated land

10_Urban area	10	8 bit integer	Pixel is classified as urban area
11_Fill value	127	8 bit integer	Out of map pixel ("Black pixels" in Fig. 2)
Land cover units	biome number	string	Land cover type ID
Land cover valid range	0-11	string	Valid range in the "Land_Cover_Type_3" dataset
Lat valid range	-9000-+9000	string	Valid range in the "Latitude" dataset
Lat/Lon fill value	25500	16 bit integer	Fill value in the "Latitude" and "Longitude" datasets
Lat/Lon scale factor	0.01	32 bit floating point	Latitude and Longitude should be multiplied by the scale factor to convert their DN values to physical values
Lat/Lon units	degree	string	Units of latitude and longitude
Lon valid range	-18000 - +18000	string	Valid range in the "Longitude" dataset
Multi land cover types presented description	description	string	Description of the "Multi_land_cover_types presented" dataset. See Sect. 6.1.4.
Multi land cover types presented fill value	-9999	string	Fill value in the "Multi_land_cover_types presented" dataset
Multi land cover types presented valid range	0-4095	string	Valid range in the "Multi_land_cover_types presented" dataset

6.1.3. Datasets. Each root group contains two data sets, "Land_Cover_Type_3" and "Multi_Land_Cover_Types_presented," as well as tile group "Geolocation" with two datasets, "Latitude" and "Longitude" (Fig. 5). Description of the datasets are given in Table 7.

Table 7: L2 Land Cover type datasets

Name of dataset	Valid range	Data type	Description
Land_Cover_Type_3	0-11	8 bit integer	Land cover type
Multi land cover types presented	0-4095	8 bit integer	Information about multiple land cover types with equal frequency. See Sect. 6.1.4
Latitude	-9000-+9000	8 bit integer	Latitude
Longitude	-18000 - +18000	8 bit integer	Longitude

6.1.4. Multi land cover types presented. This dataset provides information about dominant land cover types in 10 km EPIC pixels. This information is stored in 16-bit integer number. Bits 0 to 10 represent land cover type (Fig. 6), with bit value 1 indicating dominant land cover type. For example, if water, B2 (Shrubs), B4 (Savannas) and B5 (Evergreen Broadleaf Forests) represent 10%, 30%, 30% and 30% of the pixel area, then Multi_Land_Cover Types_Presented is '110100'=52. Bit 11 with bit value 1 indicates fill value, i.e., land cover type was not identified. In this case Multi_Land_Cover Types_Presented is '100000000000'=4095.

Bit 15	Bit 14	bit 13	bit 12	bit 11	bit 10	bit 9	bit 8	bit 7	bit 6	bit 5	bit 4	bit 3	bit 2	bit 1	bit 0
0	0	0	0	Fill value	urban	Non-vegetated	B8	B7	B6	B5	B4	B3	B2	B1	water

Figure 6. Structure of Multi_Land_Cover Types_Presented

6.2. Distribution of land cover types.

Distribution of land cover types within 10 km EPIC pixel is defined as

$$LC_i = 100\% \frac{N_i}{N}. \qquad (5)$$

Here i (i=0,1,2,…,10) represents land cover type ID, N_i is the number of the ith land cover type and N is the total number of pixels; $\sum_0^{10} LC_i = 100\%$. Sets of the MODIS 500 m land cover type 3 product within 10 km EPIC pixel was used to derive this distribution.

DSCOVR EPIC land cover product file name is DSCOVR_EPIC_ANC_LCDIST_MCD12Q1_51.h5.

6.2.1. HDF file structure. In the HDF5 file, data are grouped by tiles. Each group contains distribution of land cover type within 10 km EPIC pixel. The root group directory contains a set of attributes that describes the content of the HDF5 file. Figure 7 illustrates a snapshot of the HDFView layout of SCOVR_EPIC_LCDIST_MCD12Q1_51.h5 file.

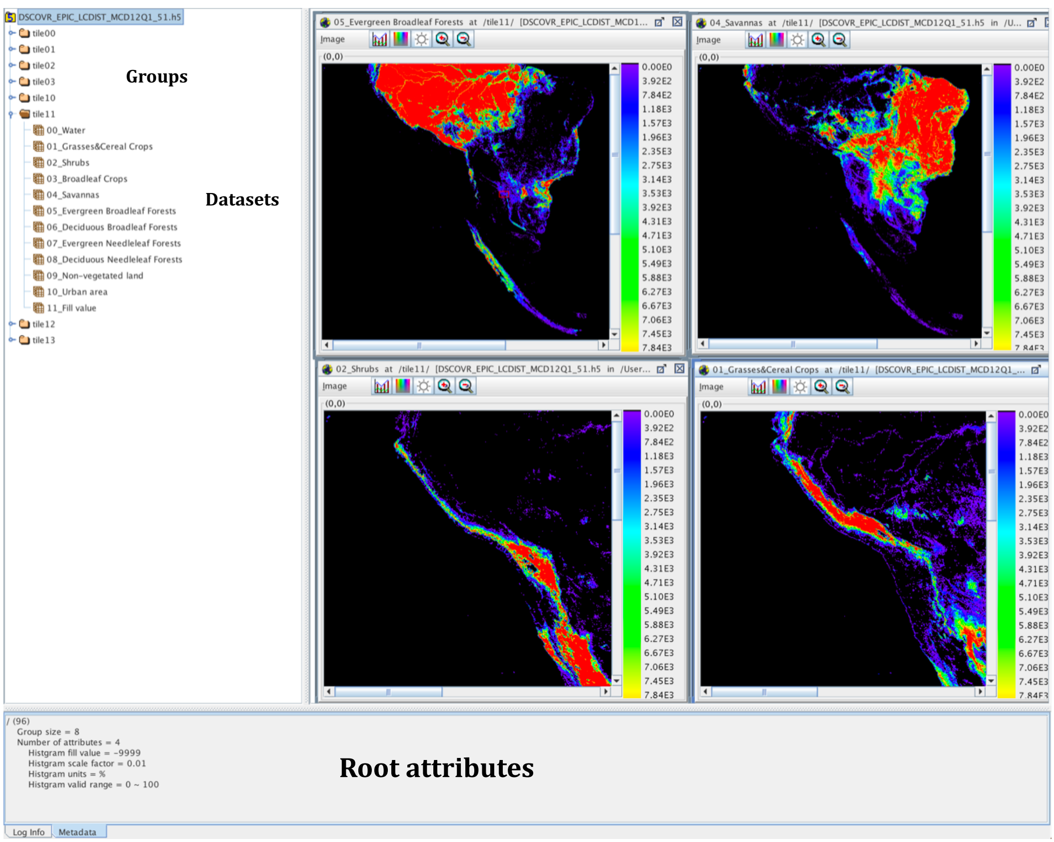

Figure 7. Structure of the ancillary 10 km SIN EPIC Land Cover Type Distribution

6.2.2. Root attributes. Root attributes include fill value, scale factor, units and distribution valid range. Details are summarized in Table 8.

Table 8: Land cover type distribution root attributes

Attribute name	Value	Type	Description
Histogram fill value	-9999	16 bit integer	Distribution fill value
Histogram scale factor	0.01	32 bit floating point	Distribution dataset should be multiplied by the scale factor to convert its DN value to physical value
Histogram units	%	string	Units of the distribution
Histogram valid range	0-100	string	Valid range after multiplying the distribution by the scale factor

6.2.3. Datasets. Each group contains distribution of land cover type. Details are summarized in Table 9.

Table 9: Land cover type distribution datasets

Name of dataset	Value	Type	Description
00_Water	0-10000	16 bit integer	Percentage of water in EPIC pixel
01_Grasses/Cereal Crops	0-10000	16 bit integer	Percentage of B1: Grasses and Cereal Crops in EPIC pixel
02_Shrubs	0-10000	16 bit integer	Percentage of B2: Shrubs in EPIC pixel
03_Broadleaf Crops	0-10000	16 bit integer	Percentage of B3: Broadleaf Crops in EPIC pixel
04_Savannas	0-10000	16 bit integer	Percentage of B4: Savannas in EPIC pixel
05_Evergreen Broadleaf Forests	0-10000	16 bit integer	Percentage of B5: Evergreen Broadleaf Forests in EPIC pixel
06_Deciduous Broadleaf Forests	0-10000	16 bit integer	Percentage of B6 : Deciduous Broadleaf Forests in EPIC pixel
07_Evergreen Needle Leaf Forests	0-10000	16 bit integer	Percentage of B7: Evergreen Needle Leaf Forests in EPIC pixel
08_ Deciduous Needle Leaf Forests	0-10000	16 bit integer	Percentage of B8: Deciduous Needle Leaf Forests in EPIC pixel
09_Non-vegetated land	0-10000	16 bit integer	Percentage of non-vegetated land in EPIC pixel
10_Urban area	0-10000	16 bit integer	Percentage of urban area in EPIC pixel
11_Fill value	0-10000	16 bit integer	Percentage of fill value in EPIC pixel

7. KNOWN ISSUES

There are still residual issues in Level 1B data that affect the geolocation accuracy. This includes errors with the star-tracker pointing, accuracy of the telescope optical model, image time stamps, and effects of atmospheric refraction. Geolocation uncertainties impact both atmospherically corrected surface reflectance and the VESDR products. Work is currently underway that treats these additional corrections to further improve science products beyond the basic requirements.

8. EXAMPLES

1. Obtaining new information on vegetation properties from the VESDR product. LAI can vary significantly with SLAI unaltered. This happens because the amount of shaded leaves can increase with SLAI unchanged. The goal of this section is to obtain canopy interceptance, i_0, direct transmittance, t_0, Clumping Index (CI) and Fractional Vegetation Cover (FVC) from LAI, SLAI and SZA.

In spite of a weak correlation between LAI and SLAI due to the above feature, these parameters satisfy the following equation (publ. [1] in sect.1.3)

$$SF = \frac{1 - \exp(-\tau)}{\tau}. \qquad \text{(E1)}$$

Here SF=SLAI/LAI is the Sunlit Fraction (SF) of leaf area, $\tau = G \cdot LAI \cdot CI/\mu$ represents the optical path through the vegetation layer, G and CI signify the geometry factor and clumping Index (CI), and $\mu = \cos SZA$. The exponent $\exp(-\tau)$ is the direct transmittance t_0. This feature allows us to derive CI and FVC using LAI, SLAI and SZA.

Problem 1: find clumping index (CI) given LAI, SLAI and SZA. This can be done based on the following algorithm

- For a given pixel calculate SF= SLAI/LAI;
- Solve Eq. (E1) for τ;
- Estimate CI as $CI = 2\tau\mu/LAI$.

The following feature of the VESDR algorithm is used here. The algorithm accumulates LAI and corresponding direct transmittances for which simulated and observed BRFs agree within model and observation uncertainties. Mean LAI is reported as retrieved LAI. The algorithm averages direct transmittance, $\bar{t}_0 = \text{mean}\, t_0$, and calculates SF as $SF = (1 - \bar{t}_0)/|\ln \bar{t}_0|$. The SLAI is estimated as $SF \cdot LAI$. Thus, the physically based Eq. (E1) is not violated in the algorithm's framework. This trick minimizes impact of errors in LAI on SLAI. Indeed, if the retrieved $LAI = kLAI_{true}$, then an error in SF is inversely proportional to k, i.e., $SF{\sim}SLAI/kLAI{\sim}SF_{true}/k$. The product $SF \cdot LAI$ therefore tends to cancel the error k.

In the above algorithm one solves Eq. (E1) using $SF{\sim}SF_{true}/k$. Its solution is $\tau{\sim}\tau_{true}k$. The ratio $\tau/LAI{\sim}\tau_{true}k/\, kLAI_{true}$ therefore also tends to minimize impact of variation in LAI on CI.

Problem 2: Find canopy interceptance, i_0, and direct transmittance, t_0, given τ. It directly follows from Eq. (E1) that i_0 is just the product between τ and SF, i.e., $i_0 = SF \cdot \tau$. The direct transmittance is $t_0 = 1 - i_0$.

Problem 3: Find FVC given t_0. Obviously, $FVC = 1 - t_0^{\mu}$.

VESDR parameters for 2016-08-23 are used to illustrate the above algorithms. There are 20 images acquired during this day. We select biome 5 (Evergreen Broadleaf Forests) in tile11, which represent Amazonian rainforests. EPIC sees this forest within a GMT time interval between 12h and 18h. Figure E1 shows LAI images of Biome 5 in the tile11.

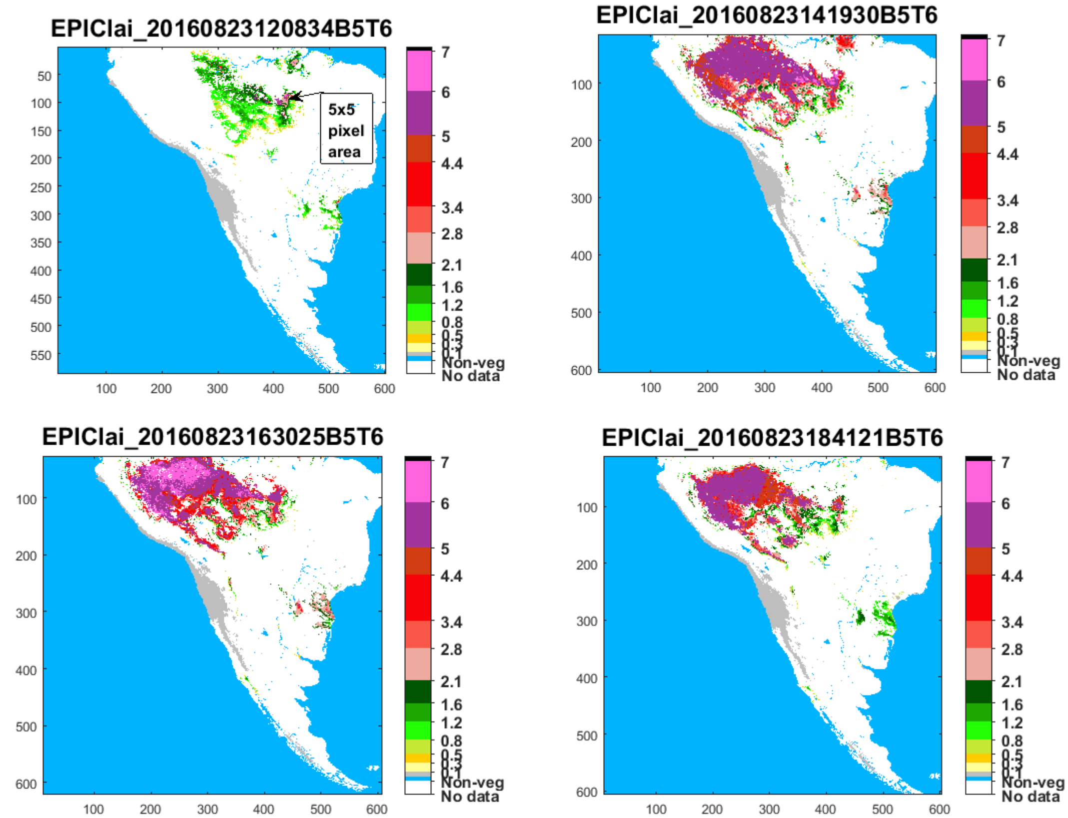

Figure E1. EPIC LAI for GMT=12,14, 16 and 18GMT. South America located at the left edge of the EPIC image at 12:08GMT. SZA for most of the pixels is about 65⁰ at this time (Fig. E2). Geo-registration error is very high, which significantly impact quality of MAIAC BRF data and consequently LAI retrievals. We do not recommend using provisional data for SZA >55º. We selected a 5x5 pixel area in Amazon, which is located between tile11 rows 97 and 101 and columns between 401 and 411, i.e., area=/tile11/01_LAI(97:101,407:411). Mean SZA in this area at 12:08GMT was 56.88º (std=0.126º)

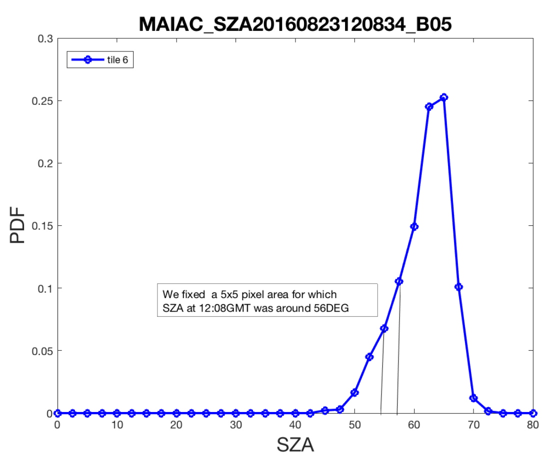

Figure E2. Distribution of SZA for B5 in the tile11 at 12:08GMT. We fixed a 5x5 pixel area for which SZA was around 56º. Mean SZA=56.88º, STD=0.126º, Coefficient of variation, 100%STD/Mean, is 0.2%

Table E1 shows mean SZA, LAI and SLAI over the 5x5 pixel area. Corresponding coefficients of variation (100%STD/Mean) are shown in parentheses. LAI should not vary within this area during a day. This is not the case in our example: variation in LAI is about 14% (or within about 0.7LAI units), which is comparable with C6 MODIS LAI (0.69 LAI units). This error is due model and observation uncertainties.

At a given GMT (i.e., for a fixed SZA), SLAI should not vary within our area. Variation in this parameter in the selected area at given SZA is about twice lower than variation in LAI (e.g., 6.6% in SLAI vs 13.1% in LAI). This confirms our theoretical results that the algorithm tends to minimize impact of errors in LAI on SLAI retrievals.

CI and FVC should be constant within our area. Their variations are about 3.1% and 3.7% (cf. with 14% in LAI), respectively, which can be treated as the precision of these parameters. This is a good value for the product precision.

Table E1. Mean LAI, SLAI, τ, CI, t_0 and FVC over our 5x5 pixel area. Coefficients of variation, Var=100%std/mean, are shown in parentheses. Last three columns show daily mean, std and coefficient of variation for SZA independent variables.

GMT,hh:mm	***12:08***	***13:14***	***14:19***	***15:24***	***16:30***	***17:35***	***18:41***	***Daily mean***	***STD***	***Var, %***
SZA (var,%)	56.88 (0.2)	42.15 (0.4)	29.03 (0.5)	20.70 (0.6)	23.28 (0.6)	34.15 (0.3)	48.15 (0.3)	n/a	n/a	n/a
LAI (var,%)	5.32 (13.1)	5.99 (3.9)	6.02 (1.9)	6.34 (2.8)	5.59 (7.6)	5.38 (8.0)	3.99 (30.6)	5.52	0.77	14.0
SLAI (var, %)	1.51 (6.6)	1.84 (2.1)	2.10 (0.4)	2.36 (0.0)	2.22 (2.1)	1.98 (2.5)	1.56 (12.4)	n/a	n/a	n/a
Solution of Eq. (E1), τ	3.43	3.04	2.62	2.44	2.2	2.44	2.28	n/a	n/a	n/a
Clumping Index	0.704	0.752	0.761	0.719	0.723	0.750	0.763	0.739	0.023	3.13
Direct transmittance	0.025	0.062	0.088	0.091	0.126	0.100	0.110	n/a	n/a	n/a
Fractional Vegetation Cover	0.975	0.938	0.912	0.909	0.873	0.900	0.889	0.914	0.033	3.66